



GÖTEBORGS UNIVERSITET  
Samhällskunskap

# Formativ bedömning

En experimentell studie om huruvida formativ bedömning ökar elevers lärande

Examensarbete i Samhällskunskap

Termin: HT11

Författare: Henning Setterberg

Handledare: Aksel Sundström

Antal ord: 9996

<b>Titel:</b>	Formativ bedömning
<b>Författare:</b>	Henning Setterberg
<b>Termin och år:</b>	Höstterminen 2011
<b>Institution:</b>	Samhällsvetenskapliga institutionen vid Göteborgs Universitet
<b>Handledare:</b>	Aksel Sundström
<b>Nyckelord:</b>	Bedömning, Bedömning för lärande, Formativ bedömning, Lärande

## Abstract

<b>Syfte:</b>	Syftet med denna studie är att undersöka om formativ bedömning förbättrar lärandet
<b>Frågeställning:</b>	Finns det ett samband mellan uppgiftsrelaterad formativ bedömning och ökat lärande?
<b>Bakgrund:</b>	Tidigare forskning visar att det existerar en kontrovers huruvida formativ bedömning har en positiv effekt på lärande. Skolverket menar att mer forskning behöver göras inom den svenska kontexten. Att undersöka om formativ bedömning har en positiv effekt på lärandet kan således anses vara av vetenskaplig vikt.
<b>Metod:</b>	För att få ett svar på frågeställningen har ett experiment genomförts. Elever har vid två tillfällen fått genomföra ett test. Experimentgruppen har mellan de två tillfällena fått ett stimuli i form av en formativ bedömning. Med hjälp av en statistisk analys så har resultatet analyserats för att se om stimulit har haft någon effekt på lärandet.
<b>Resultat:</b>	Studien visar inte på ett samband mellan uppgiftsrelaterad formativ bedömning och ökat lärande. Teorin kan dock inte anses vara motbevisad.

# Innehållsförteckning

## Abstract

1. Inledning .....	1
2. Teoretiska utgångspunkter och tidigare forskning .....	5
2.1 Sociokulturellt perspektiv på lärande.....	5
2.1.1 <i>Syn på kunskap</i> .....	5
2.1.2 <i>Syn på lärande</i> .....	6
2.1.3 <i>Lärarens och elevens roll i undervisningen</i> .....	6
2.1.4 <i>Syn på bedömning</i> .....	7
2.2 Det målrelaterade betygssystemet .....	8
2.3 Bedömning – Formativ bedömning .....	8
2.3.1 <i>Bedömning kopplat till lärande</i> .....	8
2.3.2 <i>Bedömning för lärande respektive bedömning av lärande</i> .....	9
2.3.3 <i>Formativ bedömning – En övergripande definition</i> .....	9
2.3.4 <i>Formativ bedömning – Olika aspekter, olika delar</i> .....	10
2.3.5 <i>Formativ bedömning – Gemensamma aspekter, gemensamma delar</i> .....	11
2.3.6 <i>Formativ bedömning – Olika typer av återkoppling (feedback)</i> .....	12
2.3.7 <i>Formativ bedömning – Tidsaspekt</i> .....	12
3. Precisering av syfte och frågeställning .....	13
3.1 Frågeställning .....	13
4. Design, metod och material.....	14
4.1 Experimentell design.....	14
4.2 Studiens design .....	16
4.3 Alternativ design – Alternativ metod.....	16
4.4 Urval.....	17
4.5 Operationalisering av stimuli (formativ bedömning) .....	19
4.6 Mätinstrument och stimuli .....	21
4.7 Kriterier för slutsatsdragning .....	21

4.8 Analysmetod .....	22
4.9 Validitet.....	22
4.9.1 <i>Externa förklaringsfaktorer</i> .....	22
4.9.2 <i>Intern validitet</i> .....	23
4.9.3 <i>Extern validitet</i> .....	24
4.10 Etisk hänsyn .....	25
5. Resultat .....	27
5.1 Resultatredovisning – Randomisering .....	27
5.2 Resultatredovisning – Medelvärdesresultat vid mätpunkt ett .....	28
5.3 Resultatredovisning – Medelvärdesresultat vid mätpunkt två .....	29
5.4 Resultatredovisning – Bortfallsanalys .....	30
5.5 Resultatredovisning – Jämförelse mellan de två grupperna, skillnaden från mätpunkt ett till mätpunkt två .....	30
6. Diskussion och slutsatser .....	32
6.1 Studiens resultat .....	32
6.2 Undersökningen i relation till syftet och tidigare forskning .....	32
6.3 Studiens bidrag och vidare forskning .....	32
6.4 Relevans för professionen .....	33

## Referenser

## BILAGA 1

## BILAGA 2

# 1. Inledning

## – Vilken är den viktigaste skolfrågan?

Det är ju att vi höjer kvalitén, resultaten i den svenska skolan. Och ska man lyckas göra det så är det en lång rad förändringar som man måste göra. Vi måste ha högre krav i lärarutbildningen, vi behöver ha betyg tidigare, vi behöver ha tydligare mål och krav i kursplaner till exempel (Youtube, 2011).

Det är utbildningsminister Jan Björklunds svar till lärarförbundet i ett reportage angående skolpolitiken i Sverige. I Sverige har skoldebatten gått varm i och med alliansens intåg i regeringen och huvudrubrikerna har lytt att ”skolan måste förändras”. I fokus har frågor kring *betyg och bedömning* stått och varit ett centralt tema för den skolpolitiska debatten. Frågorna har handlat om olika typer av bedömningssystem, kunskapssyn, standardiserade test (ex nationella prov) och betyg i tidigare åldrar (såsom för årskurs 6), samt vilka effekter dessa får på lärandet. Bedömningsdebattens kärnfrågor bottnar i skiljelinjen mellan det som i forskningen kallas för *bedömning för lärandet* (eng. *assessment for learning*) respektive *bedömning av lärandet* (eng. *assessment of learning*) (Gerrevall, 2008, s.93, 96; Wiliam, 2010, s.23-24; Lundahl, 2010, s.287-290; Skolverket e, 2011). I den svenska kontexten omnämns dessa två bedömningssätt vanligtvis i form av *formativ-* och *summativ bedömning*, trots att det finns en viss skillnad i begreppens innebörd, vilket jag kommer att återkomma till. Skolverket väljer att definiera formativ- respektive summativ bedömning som följer:

En bedömning med avsikten att ta reda på vad eleven lärt sig har ett summativt syfte medan en bedömning med avsikten att stärka elevens lärande har ett formativt syfte (Skolverket a, 2011).

Detta kan ses som en övergripande definition av de olika bedömningssätten och de kan sägas bottna i olika synsätt och diskurser (Gerrevall, 2008, s.93)(Skolverket f, 2011).

I Sverige anses det ha skett ett paradigmskifte i hur vi ser på bedömning och dess syfte, något som skolverket väljer att definiera som ”*det nya provparadigmet*”, där syftet har gått från att avgöra elevens kunskapsläge till att lägga mer tonvikt på att stärka elevens lärande (Skolverket c, 2011). Förändringen bottnar dels i en *förändrad kunskapssyn* och

dels i förändrade organisatoriska förutsättningar, såsom det *målrelaterade betygssystemet*. I Sverige har ett *sociokulturellt perspektiv* på lärande uppmärksammats, där Vygotskijs tankar kring den *proximala utvecklingszonen* och Bakhtins tankar kring det *dialogiska klassrummet* varit stora inspirationskällor (Dysthe, 2003, s.51, 65-67; Säljö, 2000, s.122; Säljö, 2010, s.191). Denna förändrade kunskapssyn är intimt förknippad med hur vi idag ser på betyg och bedömning (Lundahl, 2010, s.287). Förändringen från det *relativa betygssystemet* till det *målrelaterade*, som vi har i Sverige, är daterat till år 1994/1995 (Jansdotter Samuelsson mfl., 2008, s.19) och beskrivs enligt Christian Lundahl (2010) såsom följer:

I kombination med att bedömningarnas sorterande funktioner betonades ned (i och med övergången från relativa betyg till målrelaterade betyg) och att elevernas lust att lära blev ett mål i sig, skedde en övergång i synen på prov och bedömning – från bedömning av lärande till bedömning för lärandet (s.287).

Således finns det stöd för ett formativt bedömningssätt i läroplanen, både i Lpo94 och i Lgr11, där eleven ska vara mer delaktig i sin lärprocess och där läraren fortlöpande ska bedöma och utvärdera elevens kunskapsutveckling (Lpo94; Lgr11, se under *punkt 2.7 Bedömning och betyg*; Skolverket b, 2011).

Tillsammans har alltså dessa två, en *förändrad kunskapssyn* och ett *förändrat betygssystem*, legat till grund för ett paradigmskifte. Även internationellt kan det antydast att det skett ett paradigmskifte inom den utbildningsvetenskapliga kärnan då fler och fler vetenskapliga artiklar publicerats som attackerar frågan om bedömning och dess effekter på lärande (se ex. Cizek, 2010; Wiliam, 2010; Sadler, 1989; Hattie & Timperley, 2007). I USA diskuterades frågan i samband med 2001 års deklaration ”*No Child Left Behind*”, som poängterade vikten av standardiserade prov, för ett rättvist – likvärdigt bedömningssystem. I praktiken så visade det sig att skolorna i och med deklarationen gick mer åt ett summativt bedömningssystem vilket har kritiserats och debatterats. Ett flertal artiklar har därmed tillkommit för att diskutera för- och nackdelar med formativ- respektive summativ bedömning – dess användningsområde, syfte och effekter på lärande.

I Sverige kan vi se tecken på en liknande utveckling såsom i USA, där fler standardiserade, nationella prov eftersträvas, vilket torde tyda på att diskussionen om bedömning och dess effekter på lärande kommer att vara ett centralt tema även fortsättningsvis inom den svenska utbildningspolitiken. År 2010 gjorde Vetenskapsrådet en kartläggning av den svenska forskningen kring bedömning och drog samma slutsats som Skolverket, att den svenska forskningen kring formativ bedömning ”inte är särskilt omfattande” (Vetenskapsrådets rapportserie, 2010, s.72; Skolverket *a*, 2011). Detta i sin tur torde poängtera vikten av att fler undersökningar borde genomföras inom detta område.

I denna uppsats har fokus lagts på den forskning som berör formativ bedömning och dess effekter på lärande. Generellt kan det sägas att den forskning som jag tagit del av dels beskriver hur de definierat formativ bedömning och dels att formativ bedömning har en positiv effekt på lärandet, om det används korrekt (se ex. Cizek, 2010; Wiliam, 2010; Sadler, 1989; Hattie & Timperley, 2007). Dessa resultat står dock inte utan kritik. En artikel från Dunn & Mulvenon (2009) beskriver hur dessa ovan nämnda forskningsresultat saknar vetenskaplig validitet, speciellt vad avser *begreppsvaliditeten*. Kritiken innebär sålunda att olika forskare har definierat formativ bedömning på olika sätt vilket då skulle sänka tillförlitligheten i resultaten. Detta grundar sig bland annat i, (egen tolkning), att formativ bedömning inte styrs direkt av hur det är utformat, utan i vilket syfte det används. Detta framkommer bland annat av att ett flertal artiklar publicerats som framhäver att summativ bedömning även kan användas formativt (ex. Harlen, 2007).

Det förekommer alltså en kontrovers inom forskningen vilket får effekt på olika nivåer, allt ifrån samhällsnivå till individnivå – från skolpolitik till lärare och elev. Sammantaget poängterar detta resonemang dels den *samhällliga relevansen* och dels den *utbildningsvetenskapliga relevansen* för denna studie. Min slutsats blir således: *Att det är av vetenskaplig vikt att få ett svar på frågan om formativ bedömning har en positiv effekt på lärande.*

Studiens bidrag kan därmed sägas vara att påvisa om formativ bedömning har en positiv effekt på lärande och därigenom fungera som en pusselbit, om än liten, i den vetenskapliga diskussionen angående bedömning och dess effekter på lärande.



## **2. Teoretiska utgångspunkter och tidigare forskning**

Som vi inledningsvis sett så bottnar det som ligger till grund för en förändrad bedömningssyn, i Sverige, dels i en förändrad kunskapsbild och dels i förändrade organisatoriska förutsättningar – såsom ett *målrelaterat betygssystem*. Jag skall därför börja med att gå igenom dessa för att sedan gå över till att ge en tydligare bild av vad som menas med formativ bedömning med hjälp utifrån tidigare forskning. Avslutningsvis så kommer mitt syfte att preciseras utifrån det teoretiska resonemanget och konkretiseras ytterligare i form av en frågeställning.

### **2.1 Sociokulturellt perspektiv på lärande**

Den kunskapssyn som känns relevant att lyfta i denna uppsats är ett sociokulturellt perspektiv på lärande. Detta beror framförallt på att det är inom detta perspektiv som formativ bedömning kommer till uttryck. Jag skall försöka gå från att ge en övergripande bild av vad som menas med ett sociokulturellt perspektiv på lärande, till att mer konkret beröra vad detta perspektiv har att säga om kunskap, lärande och bedömning, samt hur perspektivet ser på lärarens och elevens roll i undervisningen.

Det skall dock påpekas att det inte existerar en sociokulturell ”inlärningsteori”, utan endast olika riktningar, med delvis olika fokus – som har en viss gemensam kärna (Dysthe, 2003, s.32). Med risk för att väva in för mycket i vad som utgör denna gemensamma kärna, så gör jag ett försök i att återge det som jag har uppfattat som centralt i ett sociokulturellt perspektiv på lärande: Ett sociokulturellt perspektiv ser lärande som en aktiv, kommunikativ, social process som är invävd i en historisk, diskursiv, språklig kontext (Dysthe, 2003, s.41-50; Säljö 2000; Säljö, 2010, s.183-194). Trots olika betoningar hos olika forskare så är det alltså dessa element som jag har hittat som centrala delar.

#### **2.1.1 Syn på kunskap**

Vad avser syn på kunskap så bottnar ett sociokulturellt perspektiv i ett *konstruktivistiskt* synsätt (Dysthe, 2003, s.34, 41). Dysthe (2003) beskriver detta såhär:

Kunskap existerar aldrig i ett vacuum; den är alltid ”situerad”, dvs. invävd i en historisk och kulturell kontext. Detsamma gäller språket, genom vilket kunskapen kommer till uttryck (s.34).

Kunskap har alltså både en historisk och kulturell prägel, och uppstår i en social gemenskap. På detta sätt skiljer sig ett sociokulturellt perspektiv från både ett *behavioristiskt* och ett *kognitivistiskt* synsätt. Behaviorismen framhåller kunskap som något objektivt givet i likhet med *positivismen*, medan kognitivismen ser kunskap framkomma ur inre mentala processer i likhet med *rationalismen* (Säljö, 2010, s.152, 159). Att ett sociokulturellt perspektiv lyfter fram den sociala dimensionen kan således ses som unikt och utmärkande, och ligger därför till grund för att just detta perspektiv lyfts fram i denna uppsats. På detta sätt frångår ett sociokulturellt perspektiv att kunskap skulle vara ett biologiskt fenomen, och pekar istället på att kunskap skapas kommunikativt genom mening och innebörd (Dysthe, 2003, s.43).

### **2.1.2 Syn på lärande**

Utifrån ett sociokulturellt perspektiv på lärande så förutsätts det att alla människor lär sig hela tiden i alla sociala sammanhang. Frågan är därför inte *om* utan *vad* vi lär oss, vilket poängterar vikten av kontexten (Dysthe, 2003, s.42). Vygotskij skiljer på lärande och utveckling, och beskrev därför dessa som två skilda processer som samspelade med varandra på ett komplext vis (Vygotskij, 1978, s.84). Lärande och utveckling beskrevs därför ha sitt ursprung i *sociala aktiviteter* som sedan övergick till en mental utveckling. Detta mentala steg i processen kallade Vygotskij för *appropriering*, vilket alltså kan sägas innebära att den lärande bemästrade och gjorde något till sitt eget – utifrån den sociala processen (Dysthe, 2003, s.78).

### **2.1.3 Lärarens och elevens roll i undervisningen**

Som vi tidigare sett så är lärande, enligt ett sociokulturellt perspektiv, en *aktiv* social process. Dysthe (2003) beskriver Vygotskijs tankar kring hur han ser på samspelet mellan den som lär och den som lär ut såhär:

Både den som lär och den eller de som undervisar är aktiva agenter i en social samverkan, och den aktiva samverkan är enligt Vygotskij en förutsättning för att lärande och utveckling skall äga rum (s.83).

Detta förstärks av Bahktins tankar kring lärandets dialogiska process (Dysthe, 2003, s.15, 50, 108-109). Utan att gå in på djupet så poängterar Bakhtin lärandets dialogiska natur och återkoppling som en viktig del av lärprocessen (Dysthe, 2003, s. 101). Fokuset ligger således i kommunikationen och hur den går till (Dysthe, 2003, s.101).

Med avstamp i ovanstående kan det sägas att samspelet mellan lärare och elev är en aktiv, dialogisk process, som innebär hänsynstagande till båda parter där yttrande och förståelse måste sättas i perspektiv till det sociala meningsskapandet. Detta måste i sin tur sättas i relation till vad som avses läras och hur situationen skapas för att möjliggöra detta lärande.

#### **2.1.4 Syn på bedömning**

Till skillnad från andra lärandeteorier/lärandeperspektiv så ses bedömning, utifrån ett sociokulturellt perspektiv, som en integrerad del av lärandet (Dysthe, 2003, s.40). Bedömning får på detta sätt en återkopplande karaktär som syftar till att främja lärandet istället för att vara en bedömning *av* lärandet (Dysthe, 2003, s.40). Detta synsätt grundar sig, som inledningsvis omnämndes, i Vygotskijs begrepp; *den proximala utvecklingszonen* (Dysthe, 2003, s.80-83, Säljö, 2000, s.122). Vygotskij definierar vad som avses med utvecklingszon på följande sätt:

It is the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers (Vygotskij, 1978, s.86).

Slutsatsen vi kan dra utifrån detta citat är således att det viktigaste inte är att bedöma den faktiska kunskapsnivån, utan vilken potential eleven har att nå lärandemålet med rätt assistans (Dysthe, 2003, s.51, Säljö, 2000, s.122). Detta får konsekvenser på flera plan. Utifrån ett *lärarperspektiv* bör således undervisningens fokus ligga på det som är i utveckling snarare än på den kunskap som redan uppnåtts (Dysthe, 2003, s.81).

Undervisningen får således en utmanande karaktär och poängterar samtidigt vikten av att läraren bör utgå från elevens förförståelse. Utifrån ett *bedömningsperspektiv* blir uppgiften att först avgöra/kartlägga elevens nuvarande kunskapsläge – kopplat till

lärandemålet, sedan tillsammans med eleven visa/förklara vad målet är, för att avslutningsvis visa på vägar/strategier för att nå målet.

Genom detta visar Vygotskij på bedömningens integrerade del av lärprocessen, samtidigt som den säger något om lärarens och elevens roll i undervisningen.

Vi ska nu gå vidare till att se vad som menas med det *målrelaterade betygssystemet* som nu används i Sverige, för att sedan se närmare på vad formativ bedömning är och hur det kommer till uttryck med hjälp av tidigare forskning.

## **2.2 Det målrelaterade betygssystemet**

Som inledningsvis nämndes dateras det målrelaterade betygssystemet till år 1994/1995 (Jansdotter Samuelsson mfl., 2008, s.19). Mer precist så var syftet med det målrelaterade betygssystemet att undervisningens innehåll inte skulle vara specificerat utan att det istället skulle fokusera på att nå kvalitativa och generella kunskapsmål (Jansdotter Samuelsson mfl., 2008, s.29). Det målrelaterade betygssystemet sätter därigenom elevens prestationer i jämförelse med de angivna kunskapsmålen, och hur väl de uppnått dessa. Systemet ger således utrymme för ett formativt förhållningssätt där, till skillnad från ett absolut betygssystem, elevens lärande kan ses mer som en process (Jansdotter Samuelsson mfl., 2008, s.20-34).

## **2.3 Bedömning – Formativ bedömning**

Redan inledningsvis har ett flertal begrepp kring bedömning florerat och min avsikt är att reda ut dem och försätta dem i sin rätta kontext. Jag kommer därför att börja med att kort gå igenom vilken koppling bedömning har till lärande, för att sedan gå över till att närma mig begreppet *formativ bedömning* – hur det (kan) definieras, hur det används och hur det förhåller sig till lärande.

### **2.3.1 Bedömning kopplat till lärande**

Bedömning kan ses som det pedagogiska verktyg läraren har för att avgöra vad och om en elev har lärt sig ett visst innehåll (Wiliam, 2010, s.18; Cizek, 2010, s.6). Wiliam (2010) beskriver kopplingen mellan bedömning och lärande på följande vis:

Assessment is the only way that we can know whether what has been taught has been learned. In a very real sense, therefore, assessment is the bridge between learning and teaching (s.18).

Bedömning beskrivs härigenom vara en bro mellan lärande och undervisning och får därmed en central roll i undervisningen. Sammantaget med ett sociokulturellt perspektiv på lärande bildar detta en utgångspunkt för uppsatsen.

### **2.3.2 Bedömning för lärande respektive bedömning av lärande**

Av den litteratur som jag har tagit del av att döma så går det en första skiljelinje mellan det som kallas för bedömning för lärande (*eng. assessment for learning*) och bedömning av lärande (*eng. assessment of learning*) (Gerrevall, 2008, s.93, 96; Wiliam, 2010, s.23-24; Lundahl, 2010, s.287-290; Skolverket e, 2011). Gerrevall (2008) visar på att dessa har sitt ursprung i olika diskurser, vilket ibland krockar med varandra (s.93, 100). Han pekar å ena sidan på en *utvecklingsinriktad diskurs*, som han förknippar med bedömning för lärande – där eleven och dess kunskapsutveckling är i fokus, och å andra sidan på en *systemmässigt inriktad diskurs*, som sätter samhällets krav på skolan att vara en kvalitativ, mätbar och meriterande verksamhet i fokus (Gerrevall, 2008, s.93-100). Formativ bedömning ingår i den utvecklingsinriktade diskursen och likställs alltså oftast med begreppet *bedömning för lärande* (Gerrevall, 2008, s.96). Broadfoot et al. (1999) skiljer dock på begreppen och menar att *bedömning för lärande* är en lämpligare definition än *formativ bedömning* för att förklara hur bedömning kan ha en positiv effekt på lärandet, då denna definition i större bemärkelse poängterar *syftet*, snarare än *utformningen* och/eller *frekvensen* av bedömningen (s.7). I denna uppsats så används dock fortsättningsvis begreppet *formativ bedömning*, och ska tolkas som en del av *bedömning för lärande*, (likt Gerrevalls tolkning). Dock så bidrar Broadfoot et al. argumentation till att poängtera att formativ bedömning, fortsättningsvis i denna uppsats, kommer att tolkas efter dess *syfte* snarare än dess *utformning*.

### **2.3.3 Formativ bedömning – En övergripande definition**

Guskey (2010) hänvisar till att det var Benjamin S. Bloom som initierade begreppet formativ bedömning (s.107). Detta efter att Bloom genomfört undersökningar där han såg att lärarens undervisning påverkade elever på olika sätt. Han tycktes se att lite variation, från lärarens undervisningssätt, gav stor variation på elevernas resultat, vilket

ledde till slutsatsen att läraren måste anpassa sin undervisning till eleverna för att de ska nå ett bättre resultat (Guskey, 2010, s.107-108). Bloom drog slutsatsen att läraren måste använda sig av bedömning som ett verktyg för lärande för att minska variationen – då detta dels möjliggör för eleven att få återkoppling på hans/hennes kunskapsutveckling, och dels skapa underlag för läraren att utgå ifrån för att guida eleven vidare (Guskey, 2010, s.108). Bloom kallade detta förhållningssätt för ”*Mastery Learning*”, och poängterade därigenom den formativa bedömningens återkopplande karaktär, och dess fokus på framtida steg – till skillnad från summativ bedömning som framförallt fokuserar på den nuvarande kunskapsnivån (Guskey, 2010, s.111). I likhet med Bloom poängterar flera forskare att formativ bedömning fokuserar på nästa steg, vilket särskiljer formativ bedömning från andra typer av bedömningar (Sadler, 1989, s.121; Black & Wiliam, 2009, s.6; Wiliam, 2010, s. 24-26)

Sadler (1989) definierar tre centrala processer som genomsyrar formativ bedömning: 1) Vad är målet? 2) Var är jag nu? Och 3) Hur ska jag nå målet? Dessa tre processteg utgör grunden för vad en formativ bedömning är (s.121). Hattie & Timperley (2007) återkopplar till Sadlers tre processer och tillägger ytterligare definitioner till de tre stegen, vilket de beskriver på följande vis:

Effective feedback must answer three major questions asked by a teacher and/or by a student: Where am I going? (What are the goals?), How am I going? (What progress is being made toward the goal?), and Where to next? (What activities need to be undertaken to make better progress?) these questions correspond to notions of feed up, feed back, and feed forward (s.86).

Hattie & Timperley visar härigenom att formativ bedömning dels ger *information* (feed up), dels ger *återkoppling* (feed back), och dels ger *underlag* för fortsatta steg (feed forward).

#### **2.3.4 Formativ bedömning – Olika aspekter, olika delar**

Formativ bedömning kan således tjäna olika syften, varav vi kan urskilja i alla fall två aspekter – en *elevaspekt* och en *läraraspekt*. Ur en elevaspekt ger formativ bedömning information om hur eleven ligger till kunskapsmässigt i förhållande till målet och ger indikationer på vad eleven behöver fokusera mer på. Ur en läraraspekt ger formativ bedömning information om hur undervisningen bör läggas upp och vilka delar den bör

fokusera på, med hänsyn till de delar som eleverna hade svårt för. (Guskey, 2010, s.111; Wiliam, 2010, s.23, 31; Sadler, 1989, s.120). Wiliam (2010) poängterar att vem som helst kan vara agent för att ge en formativ bedömning, till och med den lärande själv (s.24-25, 31). Framförallt är det tre aktörer som han belyser som potentiella agenter och det är; *läraren*, en *kamrat*, eller *eleven* själv (Wiliam, 2010, s.31). Wiliam kopplar ihop dessa tre aktörer med Sadlers tre processer för att förklara hur formativ bedömning kan bestå av olika delar – olika strategier (Wiliam, 2010, s.31). I denna uppsats kommer det dock endast att fokuseras på den formativa bedömning som läraren ger och utesluter således de formativa bedömningar som kan tänkas ske på daglig basis av skolkamrater eller av individuella insikter.

### **2.3.5 Formativ bedömning – Gemensamma aspekter, gemensamma delar**

Gemensamt bland de olika definitionerna av formativ bedömning är att det syftar till att vara vägledande och stärka fortsatt lärande och utveckling (Gerrevall, 2008, s.96; Davidson & Feldman, 2010, s.75). Återkommande centralt tema är således den formativa bedömningens *återkopplande karaktär*, något som även kommer att vara ett centralt inslag i denna undersökning. Sadler (1989) beskriver detta såhär:

Feedback is a key element in formative assessment, and is usually defined in terms of information about how successfully something has been or is being done (s.120).

Återkoppling eller feedback utgör på detta sätt, till stor del, grunden för vad en formativ bedömning består av (Wiliam, 2010, s.9). En vanligt citerad definition i detta sammanhang är Ramaprasads (1983) definition av formativ bedömning:

Feedback is information about the gap between the actual level and the reference level of a system parameter which used to alter the gap in some way (s.4).

Definitionen poängterar återkopplingens målsträvande karaktär – att minska gapet mellan den aktuella kunskapsnivån och målet. Hattie & Timperley (2007) påpekar dock att feedback i sig inte har någon effekt, utan påverkas av hur det används och hur det tas emot (s.82). Detta är något som förstärks av både Wiliam (2010) och Kulhavy (1977) som även poängterar att feedbacken måste vara relevant för eleven (Wiliam, 2010, s.19; Kulhavy, 1977, s.220). Feedbacken kan annars uppfattas som *negativ* och ger därmed ingen positiv effekt på lärandet (Hattie & Timperley, 2007, s.98). Positiv/effektiv

feedback kräver således dels att feedbacken är korrekt utformad – (med hänsyn till eleven), och dels uppfattas på korrekt sätt – (är lättillgänglig för eleven och därmed inte alltför främmande och abstrakt).

### **2.3.6 Formativ bedömning – Olika typer av återkoppling (feedback)**

I Hattie & Timperleys studie från 2007 (s.90) beskrivs fyra olika typer av återkoppling:

- Uppgiftsrelaterad
- Processrelaterad
- Metakognitiv återkoppling
- Personlig återkoppling

Den *uppgiftsrelaterade återkopplingen* visade sig vara den vanligaste förekommande återkopplingen. I stora drag innebär den att någon visar eleven vad som var rätt/fel och vad som saknades i svaret, och fokuserar således den formativa bedömningen på en specifik uppgift (Hattie & Timperley, 2007, s.90). Den *processrelaterade återkopplingen* handlar snarare om hur eleven går tillväga för att lösa uppgiften, vilka processer som behöver användas. På detta sätt är denna typ av återkoppling mer generaliserbar för eleven och kan användas vid olika typer av uppgifter (Hattie & Timperley, 2007, s.90-91). Den *metakognitiva typen* och *personliga återkopplingen* handlar dels om att stärka elevens egen förmåga att driva sig själv och dels att ge beröm (Hattie & Timperley, 2007, s.90).

### **2.3.7 Formativ bedömning – Tidsaspekt**

Wiliam (2010) fastställer att formativ bedömning kan verka under olika tidsaspekter (s.30). Med det menar han att bedömningen kan syfta till att nå olika mål som sträcker sig under olika lång tid. Han definierar därigenom tre cykler: *Long-cykel*, *Medium-cykel* och *Short-cykel*. En Long-cykel är en formativ bedömning som sträcker sig från *fyra veckor upp till ett år*, (ex. mellan olika perioder, kvartal, semestrar eller år). En Medium-cykel sträcker sig från *en till fyra veckor* (ex. mellan olika uppgiftsenheter), och en Short-cykel sker på *daglig basis – från dag till dag eller från minut till minut* (ex. mellan och inom lektioner). (Wiliam, 2010, s.30).



### 3. Precisering av syfte och frågeställning

Slutsatsen som vi kan dra utifrån inledningen, teorigenomgången och tidigare forskning är att det dels existerar en kontrovers huruvida formativ bedömning har en positiv effekt på lärande, och dels att det existerar olika definitioner på vad formativ bedömning egentligen innebär. Som tidigare forskning har visat beror detta till stor del på att begreppet formativ bedömning innefattar många olika delar, där olika aktörer är verksamma. Fler studier behöver därför göras som tar hänsyn till att formativ bedömning är ett omfattande begrepp – studier som börjar bena i och kartlägga de effekter och samband som de olika delarna kan tänkas ha på lärande. Inledningsvis kunde vi även se att USA:s utveckling mot ett mer summativt bedömningssystem möttes av stark kritik och gav upphov till en utbildningsvetenskaplig debatt huruvida inte formativ bedömning är ett lämpligare bedömningssystem ur lärandesynpunkt. I Sverige finns tendenser till en liknande utveckling, mot ett mer summativt bedömningssystem, vilket antyder på att diskussionen kring formativ- och summativ bedömning kommer att öka i aktualitet även här. Skolverket skriver även att den svenska forskningen kring formativ bedömning ”inte är särskilt omfattande” något som också stärks av vetenskapsrådets forskningsöversikt. Denna undersökning har därför som ambition att bistå arbetet med att börja bena i och kartlägga de effekter och samband som några av delarna av formativ bedömning kan tänkas ha på lärande, för att på så vis utgöra en pusselbit, om än liten, i diskussionen kring formativ bedömning. Jag har därför försökt att reducera så många externa förklaringsfaktorer som möjligt för att kunna uttala mig just om effekter och samband. Studien kommer således att lära oss mer om formativ bedömning har någon effekt på lärande, trots att den inte kan avsluta diskussionen en gång för alla. Som vi kommer att se längre fram har jag valt ut en specifik del av formativ bedömning som berör uppgiftsrelaterad formativ bedömning, mellan lärare och elev.

#### 3.1 Frågeställning

Frågeställningen blir därför varken mer eller mindre än:

*– Finns det ett samband mellan uppgiftsrelaterad formativ bedömning och ökat lärande?*

## 4. Design, metod och material

I det här avsnittet är min ambition att visa på hur jag har gått tillväga, vilka alternativ jag valt emellan och vilka metodologiska bekymmer jag påträffat. Jag börjar därför med att gå igenom vad som menas med experimentell design, för att sedan gå över till studiens design. Vidare diskuterar jag alternativa designers och metoder, för att sedan gå över till att visa på hur urvalet gått till. Därefter kommer jag att visa hur jag har operationaliserat stimuli (formativ bedömning), vilket material jag använt mig av för att mäta resultatet, vilka kriterier jag har för slutsatsdragning, samt vilken analysmetod jag använt för att tolka resultaten.

### 4.1 Experimentell design

Experimentell design är fördelaktig då man vill kunna isolera den oberoende variabeln och testa dess effekt på den beroende variabeln (Esaiasson m.fl., 2007, s.371). Genom möjligheten att kunna isolera förklaringsfaktorn ökar den interna validiteten och gör experiment till en bra metod då man vill kunna avgöra vilken effekt en faktor har på en annan (Bryman, 2011, s.54; ). En viktig del i en experimentell design är möjligheten till *manipulation* (Bryman, 2011, s.54; Esaiasson m.fl., 2007, s.103). Manipulation innebär att man som forskare förändrar situationen i en av grupperna, (man inför den oberoende variabeln – kallat *stimuli*), medan man inte förändrar situationen i den andra gruppen. Detta möjliggör att man kan se vilka effekter just den oberoende variabeln haft på den beroende variabeln (Bryman, 2011, s.54). Experimentell design medför således ett jämförande moment – mellan experimentgrupp och kontrollgrupp.

Det finns många olika typer av experimentell design. Den typen som används i denna studie är det som kallas för *klassisk experimentell design* (Bryman, 2011, s.54; Esaiasson m.fl., 2007, s.377). Klassisk experimentell design går även under namnet *före - efter design*, då designen innebär att ett test görs *före* manipulationen och ett test görs *efter* (Esaiasson m.fl., 2007, s.377). Detta möjliggör att vi kan se vilken effekt förklaringsfaktorn haft, alltså om det har blivit någon skillnad, på den beroende variabeln (Bryman, 2011, s.55).

En av de viktigaste faktorerna för ett lyckat experiment är hur *urvalet* gått till. Urvalet i en experimentell design består framförallt av två urvalsprocesser; urval av analysenheter – såkallad *selektion*, (ofta slumpmässig selektion) (eng. random selection), och urval av vilka som ska ingå i experimentgrupp respektive kontrollgrupp – såkallad *tilldelning*, (ofta slumpmässig tilldelning) (eng. random assignment) (Esaiasson m.fl., 2007, s.375). Hur urvalet av analysenheter gått till spelar en avgörande roll för studiens *generaliserbarhet*, något som jag återkommer till under *extern validitet*. Hur urvalet av grupper, (tilldelningen), gått till påverkar om det finns systematiska skillnader mellan grupperna, något som jag återkommer till under *intern validitet*. Genom en slumpmässig tilldelning minimeras denna risk och säkerställer att grupperna är väsentligt lika (Bryman, 2011, s.59; Esaiasson m.fl., 2007, s.103-104).

Kritiken som riktas mot experimentell design och experiment som metod kan delas upp i olika delar. En del rör kritik mot själva metoden – och menar på att det är omöjligt i samhällsvetenskaplig forskning att skapa autentiska (verkliga) experiment. De konstlade miljöer och situationer som experiment medför försvagar således, enligt kritiken, dess resultat (Esaiasson m.fl., 2007, s.386). En andra del rör kritik mot operationaliseringar – *begreppsvaliditeten* – vilket påpekar att det ofta förekommer en stor skillnad mellan studiens teoretiska definitioner och de operationella indikatorerna (Bryman, 2011, s.57; Esaiasson m.fl., 2007, s.387). Detta ställer frågan om studien verkligen mäter det den avser att mäta. En tredje del rör kritik riktad mot generaliserbarheten, då det oftast inom samhällsvetenskaplig forskning, vad avser experiment, är svårt att genomföra ett ”korrekt” slumpmässigt urval, då det kan vara svårt att hitta deltagare som är villiga att ställa upp. En fjärde del är kritik som berör klassisk experimentell design, och innebär att kritik riktas mot vilken påverkan det första testet (*före-testet*) har på det andra testet (*efter-testet*) (Bryman, 2011, s.56-57). Två risker med detta är att dels så kan det första testet avslöja experimentets syfte, vilket kan påverka resultatet, och dels så kan det första testet medföra att deltagarna ”lär sig”, och kan således få ett bättre resultat andra gången (Bryman, 2011, s.56). Denna kritik går dock att avstyra genom att hänvisa till att kontrollgruppen, som befinner sig under samma förutsättningar, således också tordes ha påverkats i samma riktning, vilket då gör att resultatet inte är påverkat av detta (Bryman, 2011, s.57).

Detta var lite om experimentell design – vad det innebär, hur det används, samt kritik. Vi ska nu gå över till att se hur detta har kommit till uttryck i denna studie.

## 4.2 Studiens design

Som vi ovan sett så grundar sig alltså studiens design i tanken att jag skall kunna isolera den oberoende variabeln, som i mitt fall är ”*formativ bedömning*”. Det jag söker svara på genom designen är således vilken effekt formativ bedömning har på lärandet, eller snarare om uppgiftsrelaterad formativ bedömning kan sägas öka elevers lärande. Detta undersöker jag genom att genomföra ett *före- och eftertest*. Mätningarna kommer på så vis att bestå av två steg – en mätpunkt innan manipulationen och en mätpunkt efter. Två grupper har således slumpats fram – en experimentgrupp och en kontrollgrupp, där den ena gruppen, (experimentgruppen), har utsatts för en manipulation i form av en formativ bedömning, (*mitt stimuli*), medan kontrollgruppen har fått stå utan denna manipulation. Genom en andra mätning kan jag därmed avgöra om stimulit har haft någon effekt.

Som mätinstrument har jag använt mig av delar från två nationella prov i matematik. Deltagarna har således vid två tillfällen fått genomföra ett test, där testens resultat jämförts mellan de båda grupperna och mellan de båda tillfällena. Det är alltså utifrån detta som slutsatser har dragits angående stimulits effekt.

Vi ska nu gå vidare till att se på varför just denna design har valts och varför andra designer och tillvägagångssätt har uteslutits.

## 4.3 Alternativ design – Alternativ metod

För att kunna göra en korrekt avvägning om vilken design och metod som är mest lämplig för att kunna besvara studiens frågeställning är det viktigt att i ett första led kunna definiera vad det är för typ av studie (Esaiasson m.fl., 2007, s.99). I mitt fall kan det konstateras att det rör sig om en *empirisk studie* som har en *förklarande ambition*. Studien syftar alltså till att förklara vilken effekt formativ bedömning har på lärande, vilket gör den till en *förklarande studie*. Att en oberoende variabel, såsom formativ bedömning, finns på plats från början av studien gör att studien har en teoretisk utgångspunkt. Detta skiljer sig från klassiska *teoriutvecklande studier*, där den oberoende variabeln ofta är okänd (Esaiasson m.fl., 2007, s.123-124). Av

frågeställningen att döma så handlar det om en *teoriprövande studie*, vilket visar sig genom att teorin, om att ”*formativ bedömning ökar elevers lärande*”, ställs på prov. Studien har således, (i metodologiska termer), ett *deduktivt* angreppssätt (Bryman, 2011, s.57). Designer och tillvägagångssätt som används vid en *beskrivande studie* eller vid ett *induktivt angreppssätt* kan därmed anses olämpliga. Trots denna åtskillnad så kvarstår det dock ett antal olika möjliga designer som ofta används vid en förklarande studie med ett *deduktivt angreppssätt*. För att avgöra den experimentella designens särställning får vi gå tillbaka till att se på syftet och frågeställningen, vilket ger oss indikationer om att det handlar om att se *effekter* och *samband*, något som en experimentell design erbjuder. En experimentell design är således mest fördelaktig då den kan erbjuda egenskaper som de andra inte tillhandahåller i samma utsträckning.

Nästa steg är att avgöra vilken typ av experimentell design som är mest fördelaktig, med hänsyn till problemställningen. I mitt fall kan en *före – efter design* argumenteras för. En *endast efter- design* skulle kunna vara användbar om jag använde mig av ett verkligt fall – om jag på något sätt kunde avgöra en situation där en formativ bedömning gått till och en situation utan, och utifrån detta göra en mätning om resultatet. Problemet med detta är dels att hitta en verklig situation där bedömningarna har skiljt sig så drastiskt, och dels att bevisa att inte andra förklaringsfaktorer spelat in i någon av situationerna. I en *före – efter design* är fördelen att det som händer i ena gruppen, torde även hända i andra gruppen, vilket innebär att skillnaden dem emellan kan antydats bero på stimuli (Bryman, 2011, s.57).

#### 4.4 Urval

Som vi kunde se under experimentell design så är det framförallt två urvalsprocesser som man måste ta hänsyn till. Dels är det hur man väljer ut vilka som skall delta i experimentet, så kallat *selektion*, och dels är det hur man väljer ut vilka som ska ingå i experimentgruppen respektive kontrollgruppen, så kallad *tilldelning* (Esaiasson m.fl., 2007, s.375). För att få en god extern och intern validitet är det fördelaktigt att använda sig av ett slumpmässigt urval i de båda urvalsprocesserna (Esaiasson m.fl., 2007, s.374-375). I denna studie har det dock inte skett ett slumpmässigt urval av deltagare till studien, (problem med detta förfarande kommer att diskuteras mer under *extern*

*validitet*). Det finns flera förklaringar till detta. För att ett korrekt slumpmässigt urval skall kunna gå tillväga är det avgörande att man kan definiera populationen och upprätta en urvalsram (Bryman, 2011, s.179-180). I mitt fall var alla elever i vid bemärkelse, populationen ifråga, vilket innebar en mer eller mindre omöjlig uppgift att upprätta en korrekt urvalsram till, speciellt med hänsyn till den begränsade tids- och resursram som fanns till förfogande. Det skall även tilläggas, (*som det även togs upp under experimentell design*), att det är svårt att genomföra ett korrekt slumpmässigt urval vid en experimentell design, i samhällsvetenskaplig forskning, då de ”utvalda” måste gå med på att delta i studien (Esaiasson m.fl., 2007, s.375).

Istället har urvalet i denna studie skett i likhet med ett *bekvämlighetsurval*, där flera lärare har blivit tillfrågade om jag fick genomföra studien i deras klass, varav några har ”nappat”. Det är alltså i dessa klasser som studien är genomförd i, vilket då inte kan räknas som ett slumpmässigt urval.

På grund av etiska hänsynstaganden var det endast lärare på gymnasienivå som tillfrågades (se mer under *etisk hänsyn*). För att undvika att avslöja allt för mycket om studiens syfte till deltagarna ställdes förfrågan endast till matematiklärare, med tanken att det inte skulle vara så märkligt att genomföra ett nationellt test i matematik för deltagarna i dessa klasser. Utifrån detta förfarande fick jag chansen att genomföra studien i två klasser, båda med samhällsvetenskaplig inriktning, i samspel med deras matematiklärare. Klasserna som ingick i studien kom således från samma skola, en storstadsskola, och gick samma gymnasiala linje. Urvalet slutade i en totalsumma på 44 deltagare.

Däremot har en slumpmässig tilldelning (eng. *random assignment*) skett för att säkerställa att grupperna är väsentligt lika (Esaiasson m.fl., 2007, s.375). Detta har skett genom en slumpstalstabell. Slumptalstabellen bestod av en slumpmässig blandning av lika många ettor som nollor. Ettorna symboliserade vilka som skulle ingå i experimentgruppen och nollorna de som skulle ingå i kontrollgruppen. Eftersom teori och tidigare forskning inte visat några tecken på att några speciella faktorer såsom klass, kön och etnicitet skulle kunna påverka resultatet i någon speciell riktning, så har fokus inte inriktats mot detta. (Detta innebär inte att dessa faktorer är verkningslösa, utan

endast att ingen av dem har lyfts fram som särskilda faktorer). Slumptalstabellen har sedan jämförts med klasslistorna för att placera in deltagarna i de olika grupperna. I praktiken resulterade detta i två grupper med 22 deltagare i varje grupp; tretton kvinnor och nio män i experimentgruppen, och tolv kvinnor och tio män i kontrollgruppen. Det totala urvalsantalet är något mindre än riktlinjen, ( $N > 60$ , alltså  $n > 30$  i varje grupp), för ett samhällsvetenskapligt experiment av denna karaktär, vilket kan påverka resultatet (Djurfeldt m.fl., 2010, s.229). Vid ett litet urval krävs det i regel större skillnader mellan medelvärdena, än vid ett stort urval, för att resultaten skall vara statistiskt signifikanta (Esaiasson m.fl., 2007, s.439).

#### **4.5 Operationalisering av stimuli (formativ bedömning)**

Avsnittet kring tidigare forskning visade att formativ bedömning syftar till att vara vägledande och stärka fortsatt lärande och utveckling – och är således en *bedömning för lärande* (Gerrevall, 2008, s.96; Davidson & Feldman, 2010, s.75). Det viktigaste är således inte utformningen utan själva *syftet*, något som även Broadfoot et al. (1999) poängterade (s.7). I denna studie är detta en utgångspunkt för själva operationaliseringen av begreppet. Till grund ligger även Sadlers tre processteg: 1) *Vad är målet?* 2) *Var är jag nu?* Och 3) *Hur ska jag nå målet?* (1989, s.121). Fokuset ligger således på *det nästa steg* som den lärande ska ta för att nå målet.

Vi har konstaterat att formativ bedömning innehåller många olika delar. Studiens ambition är inte att vara heltäckande, vilket synliggörs i att studien inte undersöker alla typer av formativ bedömning. Operationaliseringen utgår således från specifika delar/tolkningar av begreppet. Det första fokuset ligger på bedömningens återkopplande karaktär. Feedback är således det nyckelbegrepp som utgör vad mitt stimuli kommer att bestå av. Ett problem är Hattie & Timperleys kritik att feedback inte ger något resultat i sig, då det måste uppfattas som något positivt, begripligt och meningsfullt för eleven (2007, s.82). Detta är tyvärr något som jag inte kommer att ha möjlighet att kontrollera för, vilket kan spela roll för *resultatvaliditeten*, och vid förståelse för resultatet. Om resultatet skulle visa sig vara negativt, så kan en möjlig förklaring således vara att eleverna inte upplevde den formativa bedömningen som meningsfull eller begriplig. På

så vis kan inte teorin om att *formativ bedömning har en positiv effekt på lärandet* avfärdas helt vid ett negativt resultat.

Hattie & Timperleys studie (2007) beskrev fyra olika typer av återkoppling, varav en var en *uppgiftsrelaterad återkoppling* (s.90). Denna typ av återkoppling kommer att ligga till grund för undersökningen, och innebär att den formativa bedömningen kommer att fokusera på en specifik uppgift. Rent konkret innebär det att eleverna genomför ett test – får en muntlig återkoppling i *vad som var rätt/fel* och *vad som saknades i svaret/hur de kunde göra det bättre* (som även är i enlighet med Sadlers tre processteg, 1989, s.121, *se ovan*). Testet är således tvunget att ha absoluta svar för att rätt och fel enkelt skall kunna avgöras. Men även absoluta svar kan ha olika lösningsvägar. I den formativa bedömningen så har en dialog upprättats med experimentdeltagarna om hur de tänkte kring varje enskild fråga. Lösningsförslagen har således utgått ifrån detta, för att vidareutvecklas i enkla lösningsmodeller. Det skall dock poängteras att dessa lösningar endast är *ett* sätt att lösa uppgifterna på, vilket kan erhållas som kritik. Att svaren är absoluta minskar dock variationen av lösningsvägar vilket talar för studien.

Att den formativa bedömningen är muntlig kan ses som en nackdel med hänsyn till studiens *intersubjektivitet*, men var helt nödvändig i och med den teoretiska bakgrunden; *att mening och innebörd skapas i dialog och samspel*. En skriftlig bedömning hade således inte tagit hänsyn till detta, och jag hade heller inte kunnat vara säker på att alla hade läst bedömningen.

Slutligen så måste en sista hänsyn tas till tidsaspekten. Wiliam (2010) definierar tre olika tidscyklar som en formativ bedömning kan verka under (s.30). I denna studie används *Medium-cykel* som är den cykel som beskrivs vara mellan uppgiftsenheter – ungefär mellan en till fyra veckor (Wiliam, 2010, s.30). Rent konkret kommer detta innebära att den formativa bedömningen sträcker sig över två uppgiftsenheter, med ungefär en veckas mellanrum.



#### 4.6 Mätinstrument och stimuli

För att mäta om formativ bedömning ökar lärandet så kommer jag att använda mig av en specifik del, ifrån två nationella prov i matematik – en ifrån år 2000 och en ifrån år 2005. Delen som jag har valt kallas ”*del BI*” och skall genomföras utan miniräknare (se bilaga 2). Utifrån denna del så har jag plockat fram 10 frågor ifrån vardera prov, som är väsentligt lika (*i det avseende att de frågor efter liknande saker*). Alla frågorna har ett absolut svar, vilket är viktigt för studien. Det som mäts är således hur väl eleverna presterar på dessa test, dels innan stimulit och dels efter. Varje fråga ger 1 poäng och ingen fråga är mer värd än någon annan. Totalpoängen på varje test är således 10 poäng. Mellan de två testen kommer både kontrollgruppen och experimentgruppen att få tillbaka resultatet från det första testet. Utöver detta kommer experimentgruppen att få en formativ bedömning på hur de genomförde det första testet. Därefter genomför de ett nytt test, (mätpunkt två). Resultatet på detta test visar om stimulit har haft en effekt.

Det finns både för- och nackdelar med att välja två olika prov vid de olika mättillfällena. Fördelen med att ha olika prov är att eleven inte känner igen de exakta frågorna vilket gör att de inte kan svara utifrån att ha memorerat svarsalternativet. Nackdelen är dock att proven måste vara så pass väsentligt lika att de inte utgör en mätskillnad i de två olika tidpunkterna. De nationella proven är bra för detta ändamål. De nationella proven syftar till att svara upp mot kursens lärandemål och är således utformade på ett likartat sätt (Skolverket *d*, 2011). Även ifall frågornas formuleringar skiljer sig något åt så är innehållet väsentligt lika.

#### 4.7 Kriterier för slutsatsdragning

Studiens slutsats handlar om huruvida ett ökat lärande kan sägas ha ägt rum i och med en uppgiftsrelaterad formativ bedömning. Det är således av vikt att definiera vad som menas med lärande i denna studie. Säljö (2010) diskuterar tänkbara vetenskapliga definitioner på lärande (s.138-140). Det han kommer fram till är att lärande som begrepp kan ha lika många innebörder som tolkningar och beror således på vilken kunskapssyn som vi väljer att se på begreppet utifrån (Säljö, 2010, s.139). I den här studien har ett sociokulturellt perspektiv på lärande lyfts fram som kunskapssyn. Detta innebär att lärande skapas genom mening och innebörd i den sociala gemenskapen

(Dysthe, 2003, s.34-43). Denna mer konstruktivistiska hållning erbjuder tyvärr inga konkreta svar på denna fråga. För att närma oss en tydligare definition får vi istället se till vilken typ av formativ bedömning som avses mätas – (en uppgiftsrelaterad).

Uppgiftens resultat styr således vår mätning av lärande. Ett ökat resultat från mätpunkt ett till mätpunkt två, måste därför tolkas som ett ökat lärande. Om en skillnad påträffas mellan gruppernas resultat vid mätpunkt två, (att experimentgruppen fått bättre resultat än kontrollgruppen), så kan det därmed sägas att stimulit haft effekt på lärandet.

Ökat lärande på grund av formativ bedömning är alltså, i denna studie, begränsat till uppgiftens resultat. På så vis säger studien endast något om uppgiftsrelaterad formativ bedömning kopplat till lärande.

## **4.8 Analysmetod**

För att kunna analysera resultaten som jag får fram från de båda mättillfällena, samt för att kontrollera att grupperna är väsentligt lika, kommer jag att göra statistiska körningar. Genom en rad olika såkallade T-test, så kommer resultaten att analyseras och jämföras. Medelvärde för de båda grupperna, från de båda mättillfällena, kommer att vara en utgångspunkt för jämförelsen.

## **4.9 Validitet**

Validitet handlar i mångt och mycket om huruvida vi mäter det som vi avser att mäta och hur väl vi kan generalisera vårt resultat. Jag har valt att dela upp validitetsdelen i tre delar. En del som diskuterar om det finns möjliga externa förklaringsfaktorer som kan komma att påverka resultatet, en del som diskuterar den interna validiteten och slutligen en del som diskuterar den externa validiteten.

### ***4.9.1 Externa förklaringsfaktorer***

En klassisk experimentell design är vald för att kunna utesluta externa förklaringsfaktorer. Trots detta så finns det en rad olika faktorer som skulle kunna påverka, exempelvis om det fanns systematiska skillnader mellan grupperna eller ett stort bortfall. Som resultatdelen kommer att visa har båda dessa faktorer kontrollerats för och visat sig vara negativa – alltså icke förekommande i denna studie. Det bortfall

som uppstod från mät punkt ett till mät punkt två visade sig inte påverka resultatet (se mer under *resultat*).

#### **4.9.2 Intern validitet**

Intern validitet utgörs framförallt av *begreppsvaliditet* och *resultatsvaliditet* (Esaïasson m.fl., 2007, s.63). *Begreppsvaliditeten* handlar om huruvida de teoretiska begreppen är operationaliserade på ett bra och rimligt sätt, och *resultatsvaliditeten* handlar om huruvida vi mäter det som vi avser att mäta – vilket är beroende av bland annat en god *begreppsvaliditet* (Bryman, 2011, s.56-59; Esaïasson m.fl., 2007, s.63; Ekengren & Hinnfors, 2006, s.79-82).

##### **4.9.2.1 Begreppsvaliditet**

I denna studie är det framförallt två begrepp som operationaliserats och det är ”*formativ bedömning*” och ”*lärande*”. Vad avser operationaliseringen av formativ bedömning försökte jag genom teori och tidigare forskning argumentera fram en lämplig definition. Det jag kom fram till var dels en allmän tolkning av begreppet formativ bedömning och dels en specifik del av formativ bedömning, kallad *uppgiftsrelaterad återkoppling*. Denna definition och utvalda del applicerades sedan på ett empiriskt material i form av delar från två nationella prov i matematik. Dessa test fick sedan utgöra mätinstrumentet för hur lärande skulle mätas.

De problem som jag ser med operationaliseringen ligger dels i kritiken som Hattie & Timperley förde fram angående huruvida den formativa bedömningen upplevdes som positiv, begriplig och meningsfull (2007, s.82), och dels angående tidsaspekten.

Eftersom jag inte hade möjlighet att kontrollera för om den formativa bedömningen upplevdes som positiv, begriplig och meningsfull, finns det en risk att jag inte lyckats göra en bra och rimlig överensstämmelse mellan de teoretiska definitionerna och den operationella indikatorn, vilket kan påverka *begreppsvaliditeten*. Detta skulle i sin tur, om så var fallet, påverka studiens *resultatsvaliditet*, alltså om jag mäter det som jag avser att mäta.

#### **4.9.2.2 Resultatsvaliditet**

*Resultatsvaliditeten* är dock inte endast beroende av begreppsvaliditeten, utan även av studiens reliabilitet. Reliabiliteten handlar om tillförlitligheten i att studien har genomförts på ett noggrant och korrekt sätt, och att den på så vis inte lider av slumpmässiga fel (Ekengren & Hinnfors, 2006, s.80). Vad avser noggrannheten så har jag genomgående dubbelkollat allt både innan och efter mätningarna vilket kan tyda på ett noggrant tillvägagångssätt. Ett problem uppstod dock under undersökningens gång angående tidsaspekten – att det inte fanns tillräckligt med tid som jag hade till förfogande med deltagarna. Detta medförde att den formativa bedömningen (mitt stimuli) fick begränsas till ca tre minuter per person, vilket också kan ha haft en effekt på resultatsvaliditeten.

Med detta sagt så vill jag ändå påstå att jag, utefter de förutsättningar som jag haft, försökt att tillgodose den interna validiteten så gott som möjligt, bland annat genom att jag argumenterade ner de teoretiska begreppen till så konkret nivå som möjligt, för att möjliggöra en god operationalisering av begreppen, och att jag försökte genomföra studien så noggrant som möjligt.

#### **4.9.3 Extern validitet**

Den externa validiteten berör frågor om huruvida resultaten är generaliserbara. Urvalet spelar här en stor roll. Att urvalet i det första ledet, *selektionen*, inte var gjort slumpmässigt, medför att deltagarnas representativitet kan anses vara begränsad. Tidigare forskning påvisar i och för sig inte att det skulle existera någon skillnad mellan vilka som lär sig utav formativ bedömning och vilka som inte gör det. I formativ bedömningssynpunkt existerar det alltså inte en tydlig skillnad mellan olika elever – åldrar och inriktningar. Snarare handlar det om hur den formativa bedömningen uppfattas. På så vis kan resultatet från denna studie ändå säga något som även berör andra elever, än just de som deltog i studien.

En andra del av den externa validiteten är huruvida studien är generaliserbar till andra fall. Här vill jag påstå att så länge den formativa bedömningen är kopplad till en specifik uppgift, med möjligheten att avgöra rätt och fel, samt att det finns en möjlig lösning – så kan studiens resultat även gälla dessa fall.

#### 4.10 Etisk hänsyn

Den etiska diskussionen grundar sig i framförallt två, olika men parallella, krav:

*Forskningskravet* och *Individskyddskravet* (Vetenskapsrådet, 2002, s.5) Dessa står inte i rak motsats till varandra men hamnar ofta i konflikt med varandra. Generellt sett går det att säga att *Forskningskravet* handlar om att forskningen skall syfta till att hålla hög kvalitet, och att tillgänglig kunskap utvecklas och fördjupas (Vetenskapsrådet, 2002, s.5). *Individskyddskravet* syftar istället till, som namnet avslöjar, att skydda individen från kränkande behandling och/eller att forskningen åsamkar individen psykisk eller fysisk skada (Vetenskapsrådet, 2002, s.5). *Individskyddskravet* är i sin tur uppdelat i fyra huvudkrav; *informationskravet*, *samtyckeskravet*, *konfidentialitetskravet* och *nyttjandekravet* (Vetenskapsrådet, 2002, s.6).

I denna uppsats har jag försökt att ta hänsyn till dessa krav så gott som möjligt och i den mån det har varit förenligt med min design. Prioritet har lagts på *individskyddskravet* istället för *forskningskravet* vilket kan ha kommit att påverka kvaliteten av studien. En första hänsyn gjordes i början av studien då urvalet av elever skulle gå tillväga. Min första tanke var att genomföra studien i grundskolans senare del, då de i denna ålder är vana vid bedömning, men har inte gjort ett gymnasieval. Att välja grundskolan skulle på så vis eventuellt öka studiens generaliserbarhet, då denna grupp inte skulle vara fördelade efter gymnasielinjeval. För studier på individer under 15år krävs det dock att samtycke från vårdnadshavare inhämtas, något som inte var möjligt att genomföra med hänsyn till studiens tidsbrist, vilket innebar att denna grupp uteslöts.

Vad avser *informationskravet* har jag valt att vara så ärlig och öppen som möjligt, utan att förstöra experimentet. En information om studiens syfte gavs därför till eleverna innan studien genomfördes, fast endast till en viss del, (se bilaga 1). Istället för att tala om hela syftet med studien gavs informationen att studien syftade till att ”*se på olika faktorer som påverkar lärandet*”. Ordet experiment undveks genomgående då detta begrepp kan skapa oro och förvirring. Efter att studien var helt genomförd redovisades hela syftet med studien. I och med detta anser jag att *informationskravet* är tillgodosett. (Vetenskapsrådet, 2002, s.7-8). Vad avser *samtyckeskravet* utgick information innan studien genomfördes att studien var frivillig att genomföra och att den när som helst

kunde avbrytas om deltagarna så önskade. På så vis kan det argumenteras för att samtyckeskravet är uppfyllt. (Vetenskapsrådet, 2002, s.9-11). Namn och personuppgifter sparades inte och omnämndes inte överhuvudtaget i studien. Deltagarna var dock tvungna att skriva namn för att det skulle vara möjligt att koppla ihop den formativa bedömningen med testet. Efter mätpunkt två raderades alla namn och kodades om till *experimentgruppsdeltagare/kontrollgruppsdeltagare*. Ingen utomstående har heller fått ta del av denna information. Detta ligger både i linje med konfidentialitetskravet och nyttjandekravet (Vetenskapsrådet, 2002, s.12-14).

Utöver dessa hänsynstaganden har även hänsyn riktats till de rekommendationer som Vetenskapsrådet satt upp, på så vis att namn, högskola, institutionstillhörighet, samt kontaktuppgifter (mejladress) fanns med på informationsbladet, (se bilaga 1), för att möjliggöra för deltagarna att ställa frågor kring studien om de så önskade. Tillgång till resultatet när studien är klar har också erbjudits.

## 5. Resultat

Resultatdelen kommer att bestå av flera olika delar, eller snarare flera olika etapper. Det är därför min avsikt att i tur och ordning redovisa för dessa, för att sedan kunna gå över till att dra slutsatser kring vad jag fått fram. Resultaten har analyserats med hjälp av ett statistikprogram, där framförallt medelvärden har jämförts. För att kontrollera om resultaten är statistiskt signifikanta, alltså att resultaten inte beror på slumpen, så har ett flertal så kallade T-test genomförts, med en statistisk tillförlitlighet på 95 % signifikansnivå. För att utesluta externa förklaringsfaktorer har randomiseringen mellan grupperna, (alltså att grupperna är väsentligt lika), testats. Utöver detta har även en bortfallsanalys genomförts från mätpunkt ett till mätpunkt två. Då studien använt sig av en klassisk experimentell design är det framförallt två medelvärdesresultat som kommer att redovisas, en för *mätpunkt ett* och en för *mätpunkt två*. Därefter kommer en jämförelse att göras mellan gruppernas resultatskillnader från mätpunkt ett till mätpunkt två, för att se om stimulit har haft en effekt.

### 5.1 Resultatredovisning – Randomisering

Som det omnämndes under externa förklaringsfaktorer är det viktigt att grupperna är väsentligt lika. Utan detta så kan andra (externa) faktorer ha påverkat resultatet. Då teori och tidigare forskning *ej* har påvisat att det föreligger en skillnad mellan de som lär sig av formativ bedömning och de som inte lär sig av formativ bedömning, eller snarare att någon specifik grupp skulle lära sig bättre av formativ bedömning – blir min slutsats att resultatet, (medelvärdesskillnaden mellan grupperna), vid *mätpunkt ett* är den avgörande måttstocken för om grupperna är väsentligt lika. Alltså om det inte förekommer en signifikant skillnad mellan gruppernas resultat vid mätpunkt ett, så kan grupperna anses vara väsentligt lika. Detta testade jag genom att genomföra ett T-test. Om resultatet visar sig vara över 0,050 (då signifikansnivån är satt till 95 %), innebär detta att det *inte* finns en signifikant skillnad mellan grupperna – alltså att grupperna är väsentligt lika.

**Tabell 1.**

*T-test: Signifikans - medelvärdesskillnader mellan grupperna vid mätpunkt ett*

<i>Antal rätt vid mätpunkt ett, kontrollgrupp/experimentgrupp</i>	0,642
---	-------

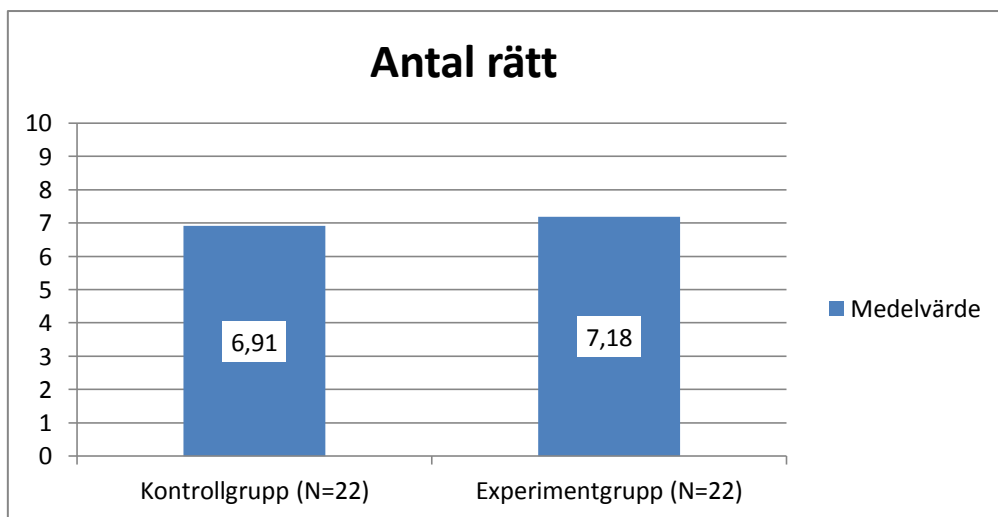
Tabellen visar att det inte finns en signifikant skillnad mellan gruppernas resultat vid mätpunkt ett, (då resultatet var över 0,050), vilket innebär att grupperna kan sägas vara väsentligt lika vad avser förkunskaper. Randomiseringen kan således antas vara lyckad.

## 5.2 Resultatredovisning – Medelvärdesresultat vid mätpunkt ett

Datamängden som framkom vid mätpunkt ett matades in i statistikprogrammet, som tog fram medelvärdet för de båda grupperna. Resultatet visas i diagrammet nedan där den beroende variabeln är antal rätt (skala 1-10).

**Diagram 1.**

*Testresultat vid mätpunkt ett*



Som vi kan se visar resultaten en liten skillnad mellan de två grupperna. Men som vi såg ovan i tabell 1 så var resultatet inte statistiskt signifikant – det är alltså med 95 % sannolikhet att skillnaden beror på slumpen.

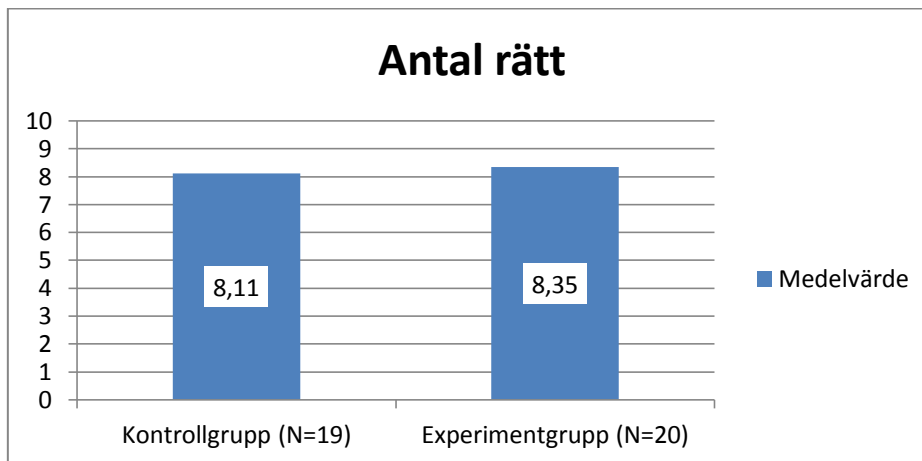


### 5.3 Resultatredovisning – Medelvärdesresultat vid mätpunkt två

Hittills har vi sett; dels att grupperna var väsentligt lika vid mätpunkt ett, och dels vilket medelvärde de båda grupperna hade vid mätpunkt ett. Detta ligger till grund för när vi nu ska se på resultaten för mätpunkt två. Vi kan räkna med en ökning i medelvärdet för de båda grupperna, såsom det omnämndes under experimentell design (att grupperna ”lär sig” från mätpunkt ett till mätpunkt två). Det intressanta är dock om det existerar en skillnad mellan de båda grupperna vid mätpunkt två. Om Experimentgruppen har ett bättre resultat än Kontrollgruppen kan stimulit antas ha haft en effekt – alltså att uppgiftsrelaterad formativ bedömning ökar lärandet. Resultatet visas i diagrammet nedan där den beroende variabeln är antal rätt (skala 1-10).

**Diagram 2.**

#### *Testresultat vid mätpunkt två*



Diagrammet visar på en viss skillnad mellan grupperna. Trots detta kan vi ännu inte avgöra om stimulit har haft en effekt. Först måste vi se dels om resultatet var signifikant, dels om bortfallet påverkat resultatet, och dels om skillnaderna mellan grupperna var signifikant. För att se om resultatet var statistiskt signifikant – alltså att det inte var slumpmässigt, genomfördes ett T-test vilket visar sig i tabellen nedan.

## Tabell 2.

*T-test: Signifikans - medelvärdesskillnader mellan grupperna vid mätpunkt två*

<i>Antal rätt vid mätpunkt två, kontrollgrupp/experimentgrupp</i>	0,632
---	-------

Tabell 2 visar att resultatet inte var statistiskt signifikant, vilket innebär att resultatet kan ha varit slumpmässigt.

### 5.4 Resultatredovisning – Bortfallsanalys

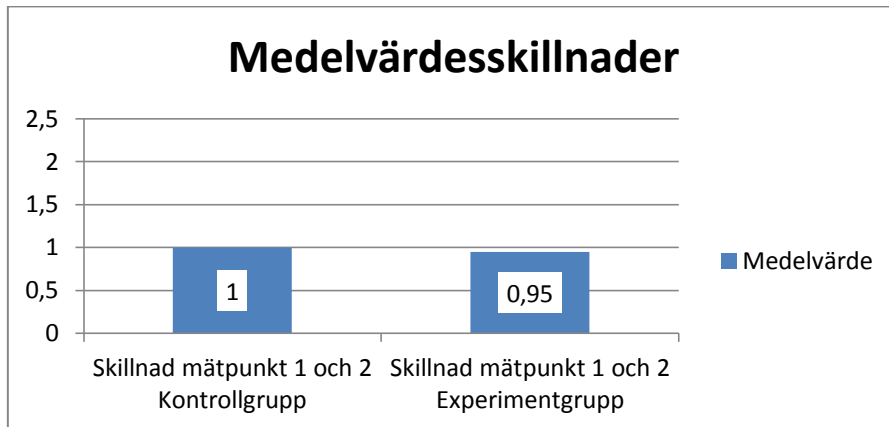
Vid mätpunkt två var det några elever som inte var där (totalt fem elever: tre i kontrollgruppen och två i experimentgruppen). För att se om detta har haft en signifikant skillnad på grupperna har jag genomfört ett T-test. Om det skulle förekomma ett stort bortfall i den ena gruppen och inte i den andra, skulle detta kunna tänkas påverka resultatet. T-testet visade att bortfallet utgjorde ca 14 % i kontrollgruppen ( $n=3/22$ ) och ca 9 % i experimentgruppen ( $n=2/22$ ). Bortfallet var alltså något större i kontrollgruppen än i experimentgruppen. T-testet visade dock att det inte fanns signifikanta skillnader mellan gruppernas bortfall, då värdet visade sig vara 0,644 – alltså över 0,050.

### 5.5 Resultatredovisning – Jämförelse mellan de två grupperna, skillnaden från mätpunkt ett till mätpunkt två

Redan vid det föregående T-testet (mätpunkt två) kan vi se att det inte existerar en markant skillnad mellan experimentgruppen och kontrollgruppen. Jag har för noggrannhetens skull även gjort ett T-test som mäter huruvida det finns en statistisk skillnad mellan gruppernas skillnad från mätpunkt ett till mätpunkt två. Alltså en jämförelse mellan experimentgruppens resultatskillnad från mätpunkt ett till mätpunkt två, med kontrollgruppens resultatskillnad från mätpunkt ett till mätpunkt två. Den visas i nedanstående diagram i form av medelvärdesskillnader och avgör frågan om stimulit har haft en effekt.

### Diagram 3.

*Testresultatsskillnader från mätpunkt ett till mätpunkt två mellan grupperna*



Diagrammet visar att det har skett en resultatmässig ökning i de båda grupperna från mätpunkt ett till mätpunkt två, vilket var förväntat. (Den observante matematikern skulle vid det här laget poängtera att medelvärdesskillnaden inte överrensstämmer med den egna subtraheringen från de båda mätpunkterna. Detta beror på att statistikprogrammet tar med bortfallet i beräkningen, och tar på så vis endast med de som var med vid båda mättillfällena). Det finns en liten skillnad mellan grupperna men som nedanstående tabell kommer att visa så är detta resultat inte signifikant på 95 % signifikansnivå.

### Tabell 3.

*T-test: Signifikans - medelvärdesskillnader mellan grupperna mätpunkt ett och två*

Medelvärdesskillnader mätpunkt 1 och 2 mellan de båda grupperna.	0,929
--	-------

Tabellen visar alltså att resultatet sannolikt var slumpmässigt och att det därmed inte kan påstås existera någon skillnad mellan de två grupperna efter mätpunkt två. Stimulit har således inte haft någon effekt.

## **6. Diskussion och slutsatser**

I detta avslutande kapitel är min avsikt att summera upp vad studien kommit fram till, hur väl jag lyckades svara upp mot min frågeställning, vad studiens bidrag blir till vidare forskning, samt vilken yrkesmässig relevans studien har.

### **6.1 Studiens resultat**

Resultatet blev ett så kallat icke-resultat då stimulit inte visade sig ha någon effekt. I och med den experimentella designen uteslöts en rad olika tänkbara externa förklaringsfaktorer, vilket torde tyda på att teorin om att uppgiftsrelaterad formativ bedömning ökar lärandet, ”motbevisades”. Dock så uppstod en rad olika komplikationer dels vad avser urvalets storlek, och dels vad avser operationaliseringen av begreppet formativ bedömning, såsom kontroll för hur den formativa bedömningen uppfattades – vilket medför att teorin ej kan sägas vara motbevisad. Likaså uppstod det problem vad avser tid per elev för den formativa bedömningen, (som blev begränsad till ca tre min per elev), vilket kan ha varit en ytterligare orsaksskälla till icke-resultatet. Detta är dock endast spekulationer. Säkert är dock att teorin *inte* fick stöd i och med denna studie.

### **6.2 Undersökningen i relation till syftet och tidigare forskning**

Syftet var att undersöka huruvida formativ bedömning ökar lärandet. Detta gjorde jag genom att bena i begreppet för att särskilja specifika delar som jag sedan testade i ett experiment. På så vis har jag genomfört det som jag åsyftade att göra. Studien återkopplade och byggde vidare på tidigare forskning som rörde formativ bedömning och bedömning för lärande, för att utgöra en pusselbit i diskussionen om formativ bedömning har en positiv effekt på lärande. Studien kan därmed bidra med vissa viktiga faktorer att tänka på för vidare forskning.

### **6.3 Studiens bidrag och vidare forskning**

Trots icke-resultatet står studien alltså inte utan lärdom. Studien visade på komplexiteten i att undersöka ett så omfattande begrepp som formativ bedömning, och poängterade vikten av att börja bena i begreppet och undersöka dess olika tänkbara delar. Utöver detta visade studien på att formativ bedömning, i alla fall vad avser uppgiftsrelaterad formativ bedömning, svårtligen får effekt på lärandet under tidsbrist,

litet urval och utan en mottagaraspekt. Vidare forskning bör således även få med dessa aspekter. Vad avser studiens problematik att kunna generalisera i och med urvalet, så bör vidare studier vara noggranna med att definiera sin urvalsram. På detta sätt kan vidare studier undvika denna generaliseringsproblematik. Större urval skulle även, som omnämnts, kunna leda till andra slutsatser då mindre medelvärdesskillnader skulle kunna antas vara statistisk signifikanta.

#### **6.4 Relevans för professionen**

Som det framkom i inledningen finns det stöd för ett formativt bedömningssätt i läroplanen, (*både i Lpo94 och i Lgr11*), vilket antyder att begreppet och dess effekter på lärande är viktiga att undersöka. Som vi även inledningsvis såg är bedömningsfrågor ett centralt tema i den skolpolitiska debatten och kommer med stor sannolikhet även att öka i aktualitet. Då skolverket även uppmanat att ytterligare studier behöver göras inom området, torde ämnet anses vara vitalt för professionen. Det är således inte endast studiens resultat och dess slutsatser som har relevans för professionen, utan även studien i sig – för att öka förståelsen kring omfattningen och innebörden av begreppet formativ bedömning.

Den praktiska implikationen studien har för läraryrket är att den kan ligga till grund för vidare diskussion kring formativ bedömning och bedömning för lärande. Som teori och tidigare forskning visade får lärarens syn på bedömning konsekvenser på flera nivåer, såsom för undervisningen och för elevens läroprocess. Studien kan på så vis ligga till grund för vidare diskussioner kring den praktiska verksamheten och hur den genomförs. Vad avser resultatet kan den inte sägas vägleda diskussionen i någon bestämd riktning, utan störst lärdom dras istället av studiens utformning och genomförande.

## Referenser

- Black, P. J. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, 21(1). s.5-31.
- Broadfoot, P. M., Daugherty, R., Gardner, J., Gipps, C. V., Harlen, W., James, M., et al. (1999). *Assessment for learning: Beyond the black box*. Cambridge, UK: University of Cambridge School of Education.
- Bryman, A. (2011). *Samhällsvetenskapliga metoder*. Andra upplagan. Malmö: Liber.
- Cizek, G. J. (2010). An introduction to formative assessment. Ingår i H.A. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment*. New York & London: Routledge.
- Davidson, J. & Feldman, J. (2010). Formative assessment applications of culminating demonstrations of mastery. Ingår i H.A. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment*. New York & London: Routledge.
- Djurfeldt, G., Larsson, R., Stjärnhagen, O. (2010). *Statistisk verktygslåda 1: samhällsvetenskaplig orsaksanalys med kvantitativa metoder*. Andra upplagan. Lund: Studentlitteratur.
- Dunn, K. E. & Mulvenon, S. W. (2009). A Critical Review of Research on Formative Assessments: The Limited Scientific Evidence of the Impact of Formative Assessments in Education. *Practical Assessment Research & Evaluation*, 14(7). s.1-11. Tillgänglig på Internet: <http://pareonline.net/getvn.asp?v=14&n=7>, hämtat 2011-11-29.
- Dythe, O. (Red.), (2003). *Dialog, samspel och lärande*. Lund: Studentlitteratur.
- Ekengren, A-M & Hinnfors, J. (2006). *Uppsatshandbok – Hur du lyckas med din uppsats*. Lund: Studentlitteratur.
- Esaiasson, P., Gilljam, M., Oscarsson, H., Wängnerud, L. (2007). *Metodpraktikan: Konsten att studera samhälle, individ och marknad*. Tredje upplagan. Stockholm: Norstedts juridik.

Gerrevall, P. (2008). Lärares professionalitet och betygsättningen – om bedömningens och betygssättningens dubbla karaktär. Ingår i C. Fritzell (Red.), *Att tolka pedagogens språk: perspektiv och diskurser*. Växjö: Växjö University Press.

Guskey, T. R. (2010). Formative assessment – The contributions of Benjamin S. Bloom. Ingår i H.A. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment*. New York & London: Routledge.

Harlen, W. (2007). *Assessment of learning*. London: Sage.

Hattie, J. & Timberley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), s.81-112.

Jansdotter S., M. & Nordgren, K. (Red.), (2008). *Betyg i teori och praktik*. Malmö: Gleerups.

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research*, 47(1), s.211-232.

Lundahl, C (2010). Bedömning – att veta vad andra vet. Ingår i U. P. Lundgren, R. Säljö, C. Liberg (Red.). *Lärare, Skola, Bildning – Grundbok för lärare*. Stockholm: Natur & Kultur.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28, s.4-13.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18. s.119-144.

Skolverket a, *Formativ bedömning – bedömning för lärande*.

[http://www.skolverket.se/skolutveckling/forskning/omraden/bedomning\\_och\\_betyg/formativ\\_bedomning](http://www.skolverket.se/skolutveckling/forskning/omraden/bedomning_och_betyg/formativ_bedomning), hämtat 2011-12-05.

Skolverket b, *Hur påverkas undervisningen av förändrade bedömningspraktiker?*

<http://www.skolverket.se/skolutveckling/forskning/tema/hur-paverkas-undervisningen-av-forandrade-bedomningspraktiker-1.157703>, hämtat 2011-12-05.

Skolverket c, *Hur ser bedömningspraxis ut i Sverige?*

<http://www.skolverket.se/skolutveckling/forskning/tema/hur-ser-bedomningspraxis-ut-i-sverige-1.157700>, hämtat 2011-12-05.

Skolverket d, *Om nationella prov.*

[http://www.skolverket.se/prov\\_och\\_bedomning/2.1100/2.1101](http://www.skolverket.se/prov_och_bedomning/2.1100/2.1101), hämtat 2011-12-12.

Skolverket e, *Vad är bedömning?*

<http://www.skolverket.se/skolutveckling/forskning/tema/vad-ar-bedomning-1.157698>, hämtat 2011-12-05.

Skolverket f, *Vilka är syftena med bedömning och betyg?*

<http://www.skolverket.se/skolutveckling/forskning/tema/vilka-ar-syftena-med-bedomning-och-betyg-1.157699>, hämtat 2011-12-05.

Säljö, R. (2000). *Lärande i praktiken – Ett sociokulturellt perspektiv*. Stockholm: Nordstedt juridik.

Säljö, R. (2010). *Den lärande människan – teoretiska traditioner*. Ingår i U. P. Lundgren, R. Säljö, C. Liberg (Red.). *Lärare, Skola, Bildning – Grundbok för lärare*. Stocholm: Natur & Kultur.

Vetenskapsrådets rapportserie (2010) *Svensk forskning om bedömning – en kartläggning*. 2:2010, Tillgänglig på Internet som Pdf:

[http://www.vr.se/download/18.6e3f84f912af3e345e78000720/2010\\_02.pdf](http://www.vr.se/download/18.6e3f84f912af3e345e78000720/2010_02.pdf), hämtat 2011-12-05.

Vetenskapsrådet (2002). *Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning*, Stockholm, Vetenskapsrådet. Tillgänglig på Internet som Pdf:

[http://www.vr.se/download/18.7f7bb63a11eb5b697f3800012802/forskningsetiska\\_principer\\_tf\\_2002.pdf](http://www.vr.se/download/18.7f7bb63a11eb5b697f3800012802/forskningsetiska_principer_tf_2002.pdf), hämtat 2011-12-05.

Vygotskij, L. S. (1978). *Mind in society – the development of higher psychological processes*. Cambridge, MA: Harvard University Press.



Wiliam, D. (2010). An integrative summary of research literature and implications for a new theory of formative assessment. Ingår i H.A. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment*. New York & London: Routledge.

## **Styrdokument**

Lpo 94 (Läroplan för det obligatoriska skolväsendet, förskoleklassen och fritidshemmet 1994) Tillgängligt på Internet som Pdf:

[http://www.skolverket.se/2.3894/publicerat/2.5006?\\_xurl=http%3A%2F%2Fwww4.skolverket.se%3A8080%2Fwtpub%2Fws%2Fskolbok%2Fwpubext%2Ftrycksak%2FRecord%3Fk%3D1069](http://www.skolverket.se/2.3894/publicerat/2.5006?_xurl=http%3A%2F%2Fwww4.skolverket.se%3A8080%2Fwtpub%2Fws%2Fskolbok%2Fwpubext%2Ftrycksak%2FRecord%3Fk%3D1069), hämtat 2011-11-29.

Lgr 11 (Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011).

Tillgängligt på internet som Pdf:

[http://www.skolverket.se/2.3894/publicerat/2.5006?\\_xurl=http%3A%2F%2Fwww4.skolverket.se%3A8080%2Fwtpub%2Fws%2Fskolbok%2Fwpubext%2Ftrycksak%2FRecord%3Fk%3D2575](http://www.skolverket.se/2.3894/publicerat/2.5006?_xurl=http%3A%2F%2Fwww4.skolverket.se%3A8080%2Fwtpub%2Fws%2Fskolbok%2Fwpubext%2Ftrycksak%2FRecord%3Fk%3D2575), hämtat 2011-11-29.

## **Webbaserat Filmklipp**

Youtube

<http://youtu.be/dwlbsIFa8C4>, hämtat 2011-11-21

## BILAGA 1

### *Informationstext till deltagarna*

Hej!

Mitt namn är Henning Setterberg och jag läser mitt sista år på Lärarprogrammet på Göteborgs Universitet. Just nu håller jag på att skriva mitt examensarbete i samhällskunskap och jag behöver Er hjälp!

Jag håller på att undersöka olika faktorer som påverkar lärandet och det jag behöver från Er är att Ni svarar på detta test. Testet är en del av ett nationelltprov i matematik A och är utan miniräknare. Det är av yttersta vikt att Ni försöker svara på frågorna så gott som det går och att Ni svarar på testets frågor ensamma. Jag ber Er därför att flytta isär bänkarna/borden så att ni kan sitta ostörda.

Testet är beräknat till att ta ca 10-15min vilket innebär att det är ett litet test, med tio frågor. Jag vill poängtera att testet är frivilligt och att Ni när som helst kan välja att avbryta, men det är till stor hjälp för mig och mitt examensarbete om ni genomför hela testet. Testet kan även vara fördelaktigt för Er på så vis att det blir en övning inför de nationella proven i maj.

Jag ber er skriva namn på förstasidan. Det är endast till för att jag skall kunna återkomma med testet till Er och kommer att raderas därefter. I studien kommer inga namn att användas och Ni förblir således helt anonym.

**JAG ÄR OERHÖRT TACKSAM FÖR ATT NI HJÄLPER MIG I MITT EXAMENSARBETE!**

Tack på förhand!

//Henning Setterberg, Göteborgs Universitet

Vid frågor eller funderingar kring studien kan jag kontaktas via e-mail:

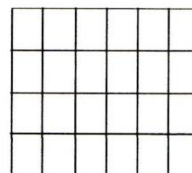
[gussetthe@student.gu.se](mailto:gussetthe@student.gu.se)

## BILAGA 2

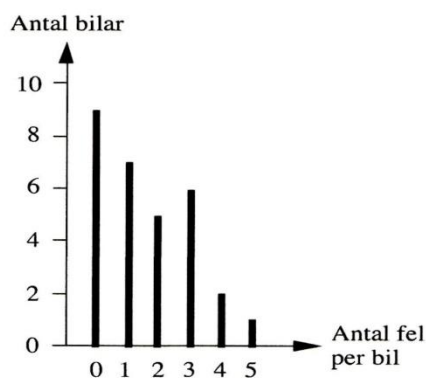
Valda delar ur nationella proven i matematik från år 2000 och år 2005

### Test 1: Valda delar ur nationella provet i matematik år 2000

1. Skugga  $\frac{5}{8}$  av figuren.



2. På en av Svensk Bilprovvnings stationer noterades antalet fel per bil under en dag. Resultatet visas i diagrammet nedan.



a) Hur många bilar undersöktes denna dag?

Svar: \_\_\_\_\_

b) Bestäm medianen för antalet fel per bil.

Svar: \_\_\_\_\_

3.  $\frac{1}{2} \cdot \frac{1}{3} =$  Ringa in ditt svar.  $\frac{2}{3}$   $\frac{2}{6}$   $\frac{1}{5}$   $\frac{1}{6}$   $\frac{2}{5}$

4. En bil kör 11 km på 10 minuter.  
Beräkna bilens medelfart i km/h.

Svar: \_\_\_\_\_ km/h

5. Lös ekvationen  $7(x - 4) = 49$ .

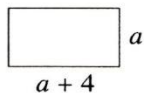
Svar:  $x =$  \_\_\_\_\_

6. Linda prismärkte alla reavaror i affären.  
Hon multiplicerade alla gamla priser med 0,85.  
Sedan skrev hon en skylt till fönstret.  
Vad skrev hon på skylten?

Svar:

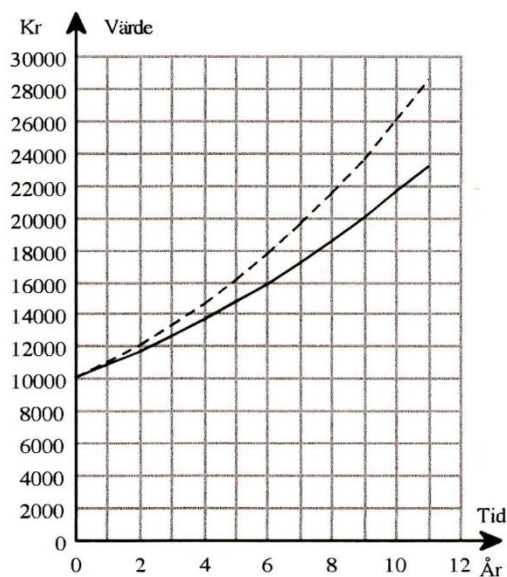
Rabatt \_\_\_\_\_ %

7. Skriv ett uttryck för nedanstående rektangels omkrets.



Svar: \_\_\_\_\_

8. Diagrammet visar hur värdet av två olika fonder växte under en period. Den årliga procentuella tillväxten var 8 % respektive 10 %.  
Besvara följande frågor utifrån diagrammet.



- a) Hur stor var fondens värde efter 10 år om den årliga procentuella tillväxten var 10 %?

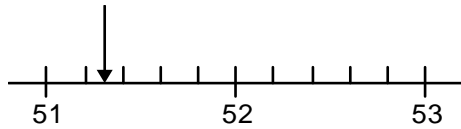
Svar: \_\_\_\_\_ kr

- b) Hur mycket längre tid krävdes för att fördubbla värdet när den årliga procentuella tillväxten var 8 % i stället för 10 %?

Svar: \_\_\_\_\_ år

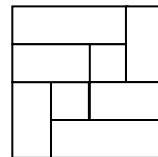
**Test 2: Valda delar ur nationella provet i matematik år 2005**

1. Vilket tal pekar pilen på?



Svar: \_\_\_\_\_

2. Skugga  $\frac{3}{8}$  av figuren.



3. Vad är 20 % av 50 kr?

Svar: \_\_\_\_\_ kr

4. Hur mycket vatten ryms ungefär i ett dricksglas?  
Ringa in ditt svar.

200 ml

200 cl

200 dl

200 hl

200 kl

5. Lös ekvationen  $7(x - 3) = 49$

Svar:  $x =$  \_\_\_\_\_

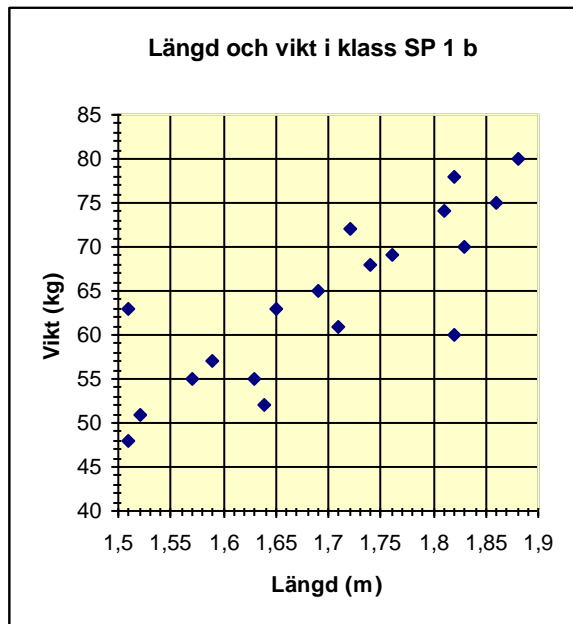
6. Placera talen **25** och **102** och **0,1** i rutorna så att resultatet blir så stort som möjligt.

	—	
<hr/>		

7. Av 6 kg äpplen får Astrid 2,8 l äppeljuice.  
Hur många liter juice kan hon få av 15 kg äpplen av samma sort?

Svar: \_\_\_\_\_ liter

8.



a) Anna går i klass SP 1 b och väger 65 kg.  
Hur lång är hon?

Svar: \_\_\_\_\_ m

b) Vilken är medianlängden i klassen?

Svar: \_\_\_\_\_ m

9. Vilket är sambandet mellan  $a$  och  $b$ ?

$a$	10	15	25	50
$b$	2	3	5	10

Svar: \_\_\_\_\_